Adaptive Behavior Models for Asymmetric Adversaries

Randy Jensen, Jeremy Ludwig
Stottler Henke Associates, Inc.
San Mateo, CA
Jensen@stottlerhenke.com
Ludwig@stottlerhenke.com

Michael Proctor, Jon Patrick
University of Central Florida
Orlando, FL
Mproctor@ucf.edu, Jonathan.patrick@us.army.mil

Wyatt Wong
Forterra Systems
San Mateo, CA
Wwong@forterrainc.com

ABSTRACT

In order for simulation based training to help prepare warfighters for modern asymmetric tactics, opponent models of behavior must become more dynamic and challenge trainees with adaptive threats consistent with those increasingly encountered by the military. In this paper we describe an adaptive behavior modeling framework designed to represent asymmetric adversaries within a multi-player virtual environment. The framework aims to provide a means for adversary models to analyze the tactical situation during execution, and adapt their behaviors and tactics accordingly. Dynamic adaptations occur both within an exercise and across exercise runs, with an automated means to carry "lessons learned" forward from one exercise to the next and adapt tactics in subsequent training sessions. This paper provides details on two distinct areas of investigation. The first area is a survey of the space of asymmetric tactics and adaptations from real-world military operations, initially focusing on urban "presence patrols". A number of training experiments were conducted in a virtual environment to solidify the behavior modeling requirements for this specific operational area, and provide a basis for generalizing to other domains. The second research area is the design and development of artificial intelligence techniques for creating adaptive adversaries. The approach makes use of an authoring tool for defining adaptive behavior models specified as partial plans that can be instantiated with choices partly driven by reward functions using data from previous events. Based on this initial behavior specification, new adaptive behaviors can be automatically generated with methods based on evolutionary algorithms. In both cases, the adversary model adapts over time in conjunction with training events.

ABOUT THE AUTHORS

Randy Jensen is a group manager at Stottler Henke Associates, Inc., working in training systems since 1993. He has developed numerous Intelligent Tutoring Systems for Stottler Henke, as well as authoring tools, simulation controls, after action review tools, and assessment logic routines. He is currently leading projects to develop automated after action review for Marine Corps combined arms training, a framework for ITS interoperability with distributed learning architectures for the Joint ADL Co-Lab, and an authoring tool for virtual training demonstrations for the Army. He holds a B.S. with honors in symbolic systems from Stanford University.

Jeremy Ludwig joined Stottler Henke in the fall of 2000 after completing his Master's Degree in Computer Science at the University of Pittsburgh with a concentration in Intelligent Systems. His research areas include behavior modeling, machine learning, and intelligent training systems.

Report Documentation Page

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
2008		**00-00-2008 to 00-00-2008**

4. TITLE AND SUBTITLE	5a. CONTRACT NUMBER
Adaptive Behavior Models for Asymmetric Adversaries	5b. GRANT NUMBER
	5c. PROGRAM ELEMENT NUMBER
6. AUTHOR(S)	5d. PROJECT NUMBER
	5e. TASK NUMBER
	5f. WORK UNIT NUMBER
7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES) **Stottler Henke Associates, Inc.,951 Mariner's Island Blvd., Suite 360,San Mateo,CA,94404**	8. PERFORMING ORGANIZATION REPORT NUMBER
9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)	10. SPONSOR/MONITOR'S ACRONYM(S)
	11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION/AVAILABILITY STATEMENT
Approved for public release; distribution unlimited

13. SUPPLEMENTARY NOTES
Interservice/Industry Training, Simulation, and Education Conference (I/ITSEC) 2008, 1-4 Dec 2008, Orlando, FL

14. ABSTRACT
In order for simulation based training to help prepare warfighters for modern asymmetric tactics, opponent models of behavior must become more dynamic and challenge trainees with adaptive threats consistent with those increasingly encountered by the military. In this paper we describe an adaptive behavior modeling framework designed to represent asymmetric adversaries within a multi-player virtual environment. The framework aims to provide a means for adversary models to analyze the tactical situation during execution, and adapt their behaviors and tactics accordingly. Dynamic adaptations occur both within an exercise and across exercise runs, with an automated means to carry ?lessons learned? forward from one exercise to the next and adapt tactics in subsequent training sessions. This paper provides details on two distinct areas of investigation. The first area is a survey of the space of asymmetric tactics and adaptations from real-world military operations, initially focusing on urban ?presence patrols?. A number of training experiments were conducted in a virtual environment to solidify the behavior modeling requirements for this specific operational area, and provide a basis for generalizing to other domains. The second research area is the design and development of artificial intelligence techniques for creating adaptive adversaries. The approach makes use of an authoring tool for defining adaptive behavior models specified as partial plans that can be instantiated with choices partly driven by reward functions using data from previous events. Based on this initial behavior specification, new adaptive behaviors can be automatically generated with methods based on evolutionary algorithms. In both cases, the adversary model adapts over time in conjunction with training events.

15. SUBJECT TERMS

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			
unclassified	**unclassified**	**unclassified**	**Same as Report (SAR)**	**11**	

Michael Proctor, LTC (Retired), Ph.D. IE, CMSP, currently is an Associate Professor with the University of Central Florida Industrial Engineering and Management Systems Department as well as with the UCF Interdisciplinary Modeling and Simulation program. His research interests include games for training, Interactive Simulation, Real-Time Simulation Agents, and Simulation-Based Life-Cycle Engineering.

Maj Jon Patrick is a Major in the United States Army Acquisition Corps and has completed his Master's Degree in Interactive Simulations and Training Systems at the University of Central Florida thru the Advanced Civil Schooling Program.

Wyatt Wong joined Forterra Systems in the fall of 2005 with over 8 years experience in design, engineering, and support of mission critical systems. Wyatt's previous experiences include areas of finance, networking, and database infrastructures. Wyatt holds a Bachelor of Applied Science in Electrical Engineering from Queen's University in Canada, as well as an MBA from the Leavey School of Business. Currently, Wyatt is focused on UI and HCI design and engineering at Forterra Systems.

Adaptive Behavior Models for Asymmetric Adversaries

Jeremy Ludwig, Randy Jensen
Stottler Henke Associates, Inc.
San Mateo, CA
Ludwig@stottlerhenke.com,
Jensen@stottlerhenke.com

Michael Proctor, Jon Patrick
University of Central Florida
Orlando, FL
Mproctor@ucf.edu, Jonathan.patrick@us.army.mil

Wyatt Wong
Forterra Systems
San Mateo, CA
Wwong@forterrainc.com

INTRODUCTION

In order for simulation based training to help prepare warfighters for modern asymmetric tactics, opponent models of behavior must become more dynamic and challenge trainees with adaptive threats consistent with those increasingly encountered by the military. In this paper we describe an adaptive behavior modeling framework designed to represent asymmetric adversaries within a multi-player virtual environment. The framework aims to provide a means for adversary models to analyze the tactical situation during execution, and adapt their behaviors and tactics accordingly. Dynamic adaptations occur both within an exercise and across exercise runs, with an automated means to carry "lessons learned" forward from one exercise to the next and adapt tactics in subsequent training sessions.

This paper provides details on two distinct areas of investigation. The first area is a survey of the space of asymmetric tactics and adaptations from real-world military operations, to generate a set of reference scenarios. A number of training experiments were conducted in a virtual environment to solidify the behavior modeling requirements for this specific operational area, and provide a basis for generalizing to other domains. The second research area is the design and development of machine learning techniques for creating adaptive adversaries. The approach makes use of an authoring tool for defining adaptive behavior models specified as partial plans that can adapt over time in conjunction with training events. This approach focuses on both supporting a natural method of encoding existing domain knowledge and the rapid adaptation of encoded behaviors. The overall objective for this approach is that the adversary behavior models should constantly challenge, and occasionally surprise, the human trainees to help them learn to be more proactive in recognizing asymmetric threats.

BACKGROUND

The training challenge is to provide the training audience with practice against an enemy who changes tactics in unpredictable or devious ways, often times specifically in response to observed patterns. Classroom lecture environments are typically not a very effective option for this training challenge, having been shown to have less than a 10% retention rate (Wiggins, 1997). While field training exercises are widely regarded as effective, they are costly and require availability of training areas and supporting infrastructure to include additional personnel to "play" the threat. Virtual simulations have typically been shown to be efficient and effective replacement for live training exercises. However, a significant limitation of current threat simulation models is their lack of dynamic asymmetric opponent behaviors reflecting recent and current adversary tactics and methods. Lacking such an opponent, training scenarios may initially create a strong positive training effect but with repeated training exercises the opponents' behavior becomes predictable and results in rapidly diminished returns on additional training.

To counter the diminishing returns of simulation-based training, opponent models of behavior must become more dynamic and contemporary. Contemporary opponent models of behavior must use asymmetric tactics, must dynamically adjust their tactics, and must generate alternative behaviors that are consistent with their perspective on warfare.

In order to provide training that exposes trainees to this kind of dynamic threat environment and the kinds of decision making they must employ in their own tactics against a thinking and reactive enemy, scripted adversary behaviors inherently cannot provide sufficient complexity to test weaknesses in the trainee's tactics. This is the motivation for the development of a system that can generate adaptive

adversary behaviors for execution in simulation based training.

TRAINING SCENARIO DEVELOPMENT

A key part of this research in laying the groundwork for the application of an adaptive behavior approach was to consider specific requirements for likely scenarios that would be used as an instrument of a training methodology. In pursuit of this goal, we developed a sample set of training scenario instances with a sequence of changing adversary tactics, in order to address the overall requirement for adaptive and asymmetric nature of insurgent tactics.

Scenario Design Objectives

The scenario content was developed to properly capture the common tactics, techniques, and procedures (TTPs) of insurgent behavior in current operational settings around the world. It was important that these TTPs not be country specific. We were aiming for a sequence of events that demonstrated adaptive insurgent tactics that consider previous successes or failures in the process of generating a response. We also were conscious of the counterbalancing effect of the simultaneous goals of realism and tractable modeling challenges. On the one hand, realism is a critical component of any training, and even more so when the goal is to familiarize the training audience with the kinds of tactical adaptations that are the hallmark of asymmetric warfare. On the other hand, full realism ultimately presents an intractable technical challenge, as the goal of modeling a complete range of possible human actions, reactions, and tactical adaptations would require a complete model of human cognition. For example the use of videotape recorder timers to trigger improvised explosive devices (IEDs) is an element of a tactic that cannot be generated by a system that has no model of such artifacts and their associated properties in its virtual world. The middle ground is a space where scenarios based on real world insurgent tactics can be eminently realistic.

We believe that a dynamic behavior adaptation model that implements a mechanism for deciding among choice points linked to these scenarios can therefore be successful in accomplishing the training goal of experience against such changing tactics. It is also possible to introduce unpredictability in such a framework, as an explicit factor in how choice points, and therefore tactics, are selected by the adversary models and performed. Factors driving unpredictable choices are taken in tandem with evidence of previous outcomes.

Another influence on scenario design involved consideration of the likely training objectives in a use case where such a scenario would be employed. The process of constructing a scenario ultimately involves the combination of underlying tactics applied in the scenario with the detailed set of events that may take place in either a linear or possibly branching manner. The level of granularity of the events within the scenario ideally should match the level at which the training audience is performing decision making, such that events that do not contribute to measurements of performance tied to intended training objectives can be abstracted out of the overall scenario event list. For example, if the training objectives concern identifying enemy tactics and deciding on proper counter-tactics, and if the training objectives do not involve the details of operational procedures that factor into counter-tactics, then these elements of the training experience can be simplified. As a result, the aim in our scenario development process was to focus on events which create the conditions where the training audience must make key decisions about the possible enemy tactics being employed. Finally, we were looking for scenarios with the potential to demonstrate the use of cultural assumptions and differences, as these increasingly play a role in asymmetric warfare. Through our literature review and consultations, over 100 documented attacks were analyzed to gain a better understanding of the trends and peculiarities involved.

Scenario Structure

Twenty-six scenario permutations were developed out of an underlying “presence patrol” scenario. To design and build these scenario permutations, the team reviewed current literature and training materials, as well as the expertise of four active duty officers who had recent experience from Iraq, Afghanistan, and Kosovo. Developing multiple scenarios from documented real world instances serves to illustrate the adaptive behavior of the adversary faced by US forces. Note that collecting scenarios for the purposes of developing adaptive behaviors requires a different level of detail than simply defining scripted sequences of events. The scenarios are informed by real world events, but decompose these into choice points which, in practice, may be automatically selected or triggered based on the adaptive logic that applies as a training event or sequence of training events unfolds. This is a key element in how the parallel objectives of realism in tactical methods and realism in tactical variation are captured.

Although this paper is not intended to enumerate specific tactical details collected from real world operational lessons learned, it is informative to describe the categories into which specific choice points fall, in an overall organizational scheme for decomposing the adaptive behaviors. These categories emerged from the review of anecdotal operational information, irrespective of observations about currently existing virtual models. The following are the major categories identified, with examples:

1. Delivery mechanism: multiple and perhaps nearly infinite means of adaptation, including IEDs with various placement methods, suicide bombers, snipers, and other combinations of specific delivery methods tied to individual capabilities.

2. Munitions type: variations of specific munitions choices given a delivery mechanism, such as the explosive used with an IED or suicide bomber, or the specific weapons used by a sniper.

3. Attack location: determined by the combination of delivery mechanism and avenues of approach and/or fields of fire needed to deliver the munitions on the target.

4. Number of attacks: one isolated, two coordinated, three coordinated, with/without the use of decoys, and other variations.

5. Environment: includes both the operational setting and the mission, which may take place in a market, street, check point, searches within houses or buildings, presence patrol on city streets, and other variations.

Using these categories, scenarios were developed that provided illustrated examples of known insurgent behaviors and trends. Scenarios were developed using current military doctrine, and for this effort all scenarios were focused on squad level dismounted patrols in an urban environment similar to Iraq. In the complete set of scenario instances, each instance shows the co-evolution of tactics by the adversary pair (insurgent vs. Coalition) over time. In any given scenario instance, the US forces conduct a tactical operation, the insurgent adversary performs a tactic designed to defeat or degrade the effectiveness of the US tactic, an outcome occurs, the US adapts and develops a counter tactic, a new outcome occurs, and finally the insurgents adapt a new tactic, which leads to the next scenario instance pairing. As evidenced in real world asymmetric operations lessons learned, sheer variation itself is a factor in the enemy's choice of new tactics, along with other considerations involved in responding to US counter tactics.

ADAPTIVE MODELING APPROACH

The scenarios were central to the construction of a decision making model for the adaptive adversary behaviors, by providing scope for the inputs and outputs that constrain the space of possible actions and reactions of the adaptive adversary model. Although the set of scenario instances represents a sample sequence of adaptations motivated by preceding successes and failures, the adaptive behavior model automatically generates differing tactics in an unscripted way. The behavior models described in this paper provide support for all of the adaptations identified in these scenario designs, with sequencing entirely driven by exercise events rather than a predefined ordering.

Our approach to the problem of behavior adaptation and creation for asymmetric adversaries contains two primary elements:

1. **Initial Insurgent Behaviors**: An initial set of insurgent behaviors that are created by a subject matter expert (SME). This captures the current knowledge of insurgent tactics and allows for realistic adversary performance from the first training simulation.

2. **Behavior Adaptation:** Adaptive choice points that are embedded in the initial simulation behaviors allow for adaptive behaviors. The SME specifies partial behaviors through the use of choice and reward points and then the system automatically learns which particular behavior(s) work best against the current adversary. For example, if an IED in a trash can is safely disarmed in one scenario, the adversary model might respond by combining a decoy trash can IED with a sniper in the next scenario.

This general approach is used to control behavior at two specific levels: **tactic** and **agent**. At the **tactic** level, the basic pieces of a training scenario are put in place before the scenario begins. This includes placing snipers, IEDs, ambush forces, etc. As the scenario

unfolds, the behavior at the **agent** level controls the behavior of each agent within the scenario.

In this paper, we focus on initial insurgent behaviors and behavior adaptation at the tactical level, though the results are directly applicable to the agent level as well.

Tactical Behavior Adaptation

Tactical behaviors determine the initial conditions of a training scenario. Prior to the start of the scenario, the insurgent forces and objects are put into place by running an initial behavior. Adaptive logic in the behaviors is used to learn the most effective adversary tactics. Adaptive choice points are used to setup all of the elements of an insurgent attack such as type of attack, IED concealment, insurgent/object placements, decoys, munitions, etc. Behavior adaptation works as follows:

1. Choices are made for the current instance of a training scenario, based on the values associated with each choice.

2. The training scenario is carried out by blue and red forces. The behavior of the red forces (insurgents) during the scenario can be controlled by Agent level behaviors or by human role players.

3. At the end of the scenario, a reward function is used to update the values associated with each of the choices that applied during the scenario.

There are two important notes to make regarding adaptation. First, based on our particular implementation of adaptation the system can actually select sub-optimal tactics for novice level players while still learning what the best tactics are. This allows a degree of flexibility when implementing this work as part of a training system. Second, adaptation works on both human- and computer-created behavior. That is, the choice point logic can learn to ignore ineffective behaviors regardless of who/what created them. This is a very useful functionality when combined with the automatic creation of new behaviors.

Additionally, synthetic behaviors naturally depend heavily on the virtual environment for location markup. The relationship between the environment and certain tactics can be very close. Enemy avenues of approach, enemy fire lanes, and lack of cover and concealment (exposure) are environmental conditions that, combined with tactical failures (failure to follow rules of engagement, etc.), may predicate an attack.

Implementation Details

The adversary behaviors were encoded in an existing graphical behavior modeling architecture (Fu & Houlette, 2002), where behaviors are composed of actions, predicates, and directed connecters that describe agent behavior. As an action in a behavior can reference a primitive action, or another behavior, hierarchal behavior networks can be created.

The work described in this paper makes use of an updated version of the behavior modeling architecture that incorporates the extended dynamic scripting algorithm (Ludwig & Farley, 2008). This particular learning algorithm was selected based on its ability to quickly learn to best an opponent in modern computer games and simulations (Ponsen & Spronck, 2004; Spronck et al., 2006).

The updated behavior modeling architecture introduces two additional types of action nodes to support a specialized reinforcement learning algorithm: choice points and reward points. Each choice point represents a decision, where the behavior model learns to select the best action from the available action. Learning occurs when the results of the scenario are received as feedback by corresponding reward nodes. In this manner, the behaviors can encode the range of adaptive behavior found during the training scenario development and choose an initial scenario configuration likely to surprise the human players.

Related Work

Choice points, as used in the behaviors described in this paper, are similar to the choice points found in the Hierarchy of Abstract Machine and ALisp architectures (Andre & Russell, 2002). The extended dynamic scripting algorithm (Ludwig & Farley, 2008) builds off of previous research on dynamic scripting (Spronck et al, 2006) and hierarchical dynamic scripting (Dahlbom & Niklasson, 2006; Ponsen & Spronck, 2004).

ADVERSARIAL BEHAVIORS

Building off of both the developed training scenarios and the extended dynamic scripting algorithm, we created behaviors to adaptively determine the best initial training scenario configuration. To do this we specify the possible training scenarios as a hierarchical

set of choice points, where the objective is to learn to select the scenario configurations most likely to succeed against the current players.

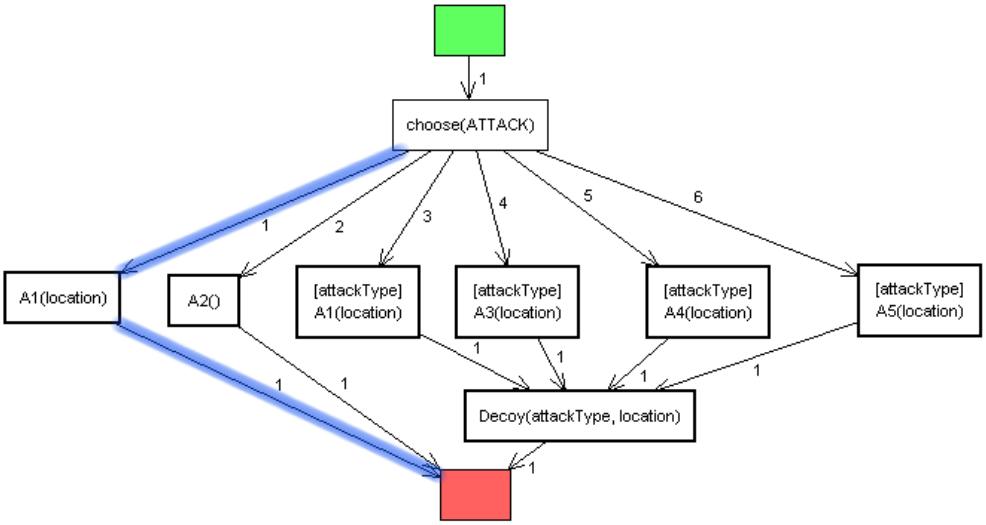

Figure 1. Main Adaptive Adversarial Behavior

An abstracted version of the top level tactical behavior is shown in Figure 1, giving example behavior structure without domain-specific content. The ATTACK choice point chooses the A1 adversarial tactic, as shown by the bold highlighting. The choice point selection is highlighted, where the A1 tactic is selected.

After the main behavior chooses the A1 attack type, control is transferred to the A1 sub-behavior, seen in Figure 2. The choice point for the A1 tactic must choose between a number of different ways to carry out the A1 tactic. In this instance, A1_2 is chosen and control transfers to the A1_2 sub-behavior.

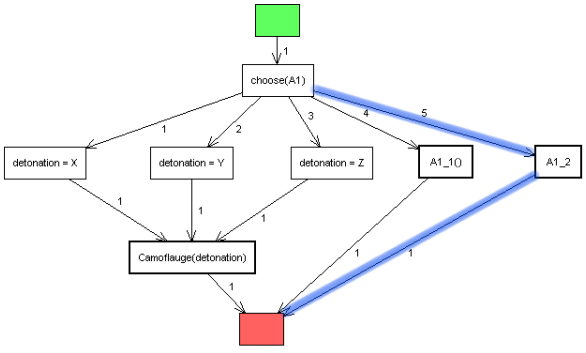

Figure 2. A1 Sub-behavior

The behavior in Figure 2 chooses the specific elements of the A1 tactic. The A1_2 behavior (Figure 3) is responsible for choosing the specific location of the A1_2 tactic in the simulated world from a number of pre-determined locations. A primitive action, selectA1_2ActiveFile, is then called to load all of the selections into the simulator world.

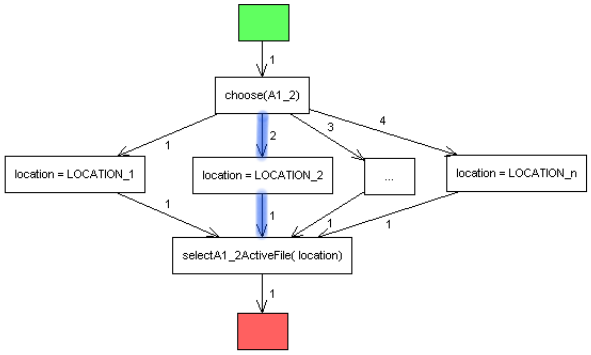

Figure 3. A1_2 Sub-behavior

The behavior in Figure 3 determines the location of the tactic elements in the simulation. After the A1_2 sub-behavior is executed, the simulation is ready for the team of trainees to begin using the simulation.

Once the simulation has ended, the results from the scenario are used to update the values associated with the actions in the choice points. This is performed with the help of a number of reward functions that supply a quantitative value that represents the success of the adversary's tactics in the scenario. For example, a reward function might take into account the current health levels of all of the members of both team or whether or not an IED was detonated. Reward points, encoded as part of the behavior model, use one or more

of these reward functions to update a choice point's action values. They may also be integrated with specific incremental or decremental reward factors that take into account considerations such as unpredictability. The research effort focused primarily on the design of the approach, as opposed to the implementation of a complete set of reward functions.

The outlined approach makes use of machine learning methods that learn, statistically, which types of tactical situations are more likely to result in a positive outcome. Some of the time, the adversary models will select the best known initial settings (exploiting) and sometimes it will search for another initial configuration that might result in even better performance (exploring). While this type of adversarial learning is not psychologically plausible, it does meet the overall training objective in that it will constantly challenge and surprise the human trainees in a quick and efficient manner.

FEASIBILITY STUDY

The objective of the feasibility study was to demonstrate what a training event would look like, with synthetic adversaries playing out the tactics generated by the adaptive behaviors. This study was carried out in the distributed and massively multi-player On-Line Interactive Virtual Environment (OLIVE) virtual environment, created by Forterra Systems ("Purpose Driven Virtual Worlds for Everyone").

The example training session was conducted with human players as participants in a sequence of scenario instances. The Reserve Officer Training Corps (ROTC) detachment at the University of Central Florida provided 7 cadets to fulfill this role. These cadets carried out a number of multi-player training sessions against adaptive insurgent tactics in the OLIVE environment. Additionally, two Arabic speaking students role-played the squad interpreter and female used in the scenarios and provided cultural context. Forterra also provided several support personnel as well located across the United States which had the added benefit of demonstrating the distributed training capability of OLIVE. After a brief familiarization with OLIVE, the training demonstrations were held over a 3 day period, totaling 9 hours of presence patrol exercises. This was carried out in OLIVE's virtual Baghdad urban terrain area.

Figure 4. Patrol Formation with Six Trainees

Figure 5. Two Trainees Taking Cover

Two screenshots from the feasibility study are shown in Figure 4 and Figure 5. In order to obtain the scenario outcome desired, several actions were planned to demonstrate common mistakes that could and have been made in typical current operating environments. Unit movement techniques, actions on contact, and common tasks were maintained in accordance with current doctrine.

As an example training event, this demonstrated that the insurgent tactics can be modeled and exercised in the OLIVE environment, and that the kinds of changing tactics that would be generated by the created adaptive behaviors can mirror authentic evolving tactics identified from the operational world. The process of collecting operational information to use as the building blocks for choice points in scenarios was conducted within a feasible scope of effort, and we similarly laid out the path for codifying these. Part of the objective in carrying out this process for a selected operational task area was to gain insight into the level of complexity for such an effort, as a reference point for potential similar work in other task domains.

CONCLUSIONS AND THE WAY FORWARD

One of the primary design goals was to yield an adaptive behavior methodology that can be applied in many different team training contexts. Therefore, as one outcome of the research, we identified some high level defining characteristics for potential training use cases in which this design may be applied:

- **Virtual Training**. The adaptive behavior design makes the general presumption that training will be conducted in a virtual environment where synthetic enemy agents can be controlled by automated behaviors applying tactics generated by the adaptive logic. The initial design makes use of the OLIVE virtual environment. It is worth noting that this may not be a universal requirement, as it is possible to imagine use cases even with live training, where the adaptive behavior mechanism could be provided with direct inputs summarizing the results of a live training event, which would then serve as the basis for a new set of tactics provided descriptively to the administrators of a subsequent live training event.

- **Tactically Oriented Domains**. This behavior modeling approach is ideally suited for training domains where there is a naturally measurable relationship between tactics and outcomes (or at least partial successes and failures) in the course of a training event. In order for a synthetic enemy to learn or adapt their methods, there must be a computer-definable notion of success or failure that can be associated with previously applied methods.

- **Scenario Oriented Domains**. The notion of gaining practice against an adversary with changing tactics naturally lends itself to a training mode organized around a framework of scenarios in which results can be evaluated and lead to adaptations between exercises. Of course for this training context the notion of a scenario is strictly a template into which variability in tactics, specific events, and other details can be enacted.

A full set of developed tactical variations can be derived from an initial set of scenarios by categorizing the choice points they rely on, and annotating the virtual environment with as many possible corresponding choice values as are feasible. This essentially results in an implemented library of possible tactical variations linked to artifacts of the virtual world. This library is capable of producing a very large number of insurgent tactics, of which the concrete scenarios are only a subset. That is, while it can perform all of the adaptations outlined in the scenarios, it is by no means limited to these scenarios.

Typical Training Use Case

For the training itself, the resulting behavior adaptation mechanism and the tactical library would remain flexible for use in a variety of contexts. A likely training use case may involve one where a small team coordinates a training event. An instructor (or potentially the team leader) provides them with their orders and they plan the operation before beginning. The exercise is conducted, and they encounter an asymmetric threat employing a certain tactic driven by the system behaviors, which either succeeds or is defeated by the training team. The exercise is concluded, and the team goes through an after action review led by the instructor (or potentially the team leader), to attempt to identify what tactics the enemy was using, how successful they were, and why. Depending on the specific goals of the instructor or team leader, they may want to make it clear exactly what the enemy tactic was, or leave it implicit in what is observable in playback. AAR playback may include review from various perspectives including the enemy perspectives, or these may be limited to only friendly force perspectives (similar to the data available from real world operations). The adaptive behavior mechanisms likewise evaluate the success or failure of enemy tactics in the exercise.

A subsequent training event is scheduled. In preparation for this, two things happen. Within the behavior mechanism, adaptations have already been developed based on the previous exercise. On the human side, the training team is required to prepare for the next operation once again, this time giving specific thought to any counter-tactics or procedures they may choose to apply in the mission. Depending on the length of the exercises, this sequence may be repeated several times, either in the same part of the virtual environment and under similar basic scenario conditions, or potentially in other areas with different buildings, cover, diversions, access points, and so on.

Instructional Methods

This adaptive framework is designed to support three instructional methods: (1) best tactics, (2) team-based tactics, and (3) dynamically adjusted tactics. In the

first case, all trainees see the "latest and greatest" behaviors as they have evolved from the entire history of training events. As such, the behaviors refer not only to specific combinations of tactical choices, but the ways in which these tactics are automatically chosen in response to the actions of the training audience. In the case of team-based tactics, each new group of trainees begins at the same starting point and the set of adversary behaviors evolve specifically in response to this particular team. For the third case, dynamically adjusted, the "latest and greatest" tactics serve as the behavior starting point; these behaviors are then automatically adjusted to make the level of play match that of the training team.

In the envisioned system, there would be one instance of the underlying behavior execution manager for each training server. This instance is responsible for defining the initial scenario setup each time a team begins a new training session and applying the resulting reward when the training session is complete. Instruction method (1) is supported by default – all training sessions started on the same server will make use of the most developed set of behaviors on that server. To support (2), we would need to add a mechanism that maps a training team to a particular behavior file. When a particular team logs on for the first time, a new copy of the behavior file is created. All adaptation/creation that occurs will only change the currently loaded behavior file. In this way, we can support team-based tactics with a small amount of additional computer programming. Finally, the underlying reinforcement learning algorithm used in the choice point mechanism has been used to successfully support dynamic difficulty adjustment (3) in previous work by Spronck et al. (2006). Transferring these results to a given domain is relatively straightforward, combining elements of both (1) and (2). In this case the system learns the best behavior (similar to 1) while simultaneously making sub-optimal choices for the particular team (similar to 2), where the goal is to allow the human team to win with some pre-set frequency (e.g. roughly 50% of the time).

While the above solution supports (1), (2), and (3) for any particular server, additional work would be required to support transferring behaviors across servers. Specifically, this would require a universal central behavior repository, where each server could download and upload behaviors. For example, in the case of (2), the behaviors for a team would be downloaded from the central repository before the training session, updated as a result of the training session, and then uploaded back to the central repository. This would allow a team to use any training server while facing the adversaries developed in response to their behavior. In the case of (1) and (3), the additional difficulty is determining which behavior is the "best" across servers, which requires methods for determining that one behavior is superior to another. In most any of these circumstances, a likely outcome is that the training audience would gain exposure to a wider variety of enemy tactics than they may otherwise encounter.

It is reasonable to anticipate that an adaptive behavior modeling framework such as this could be utilized to support individual, leader, and small unit training in many settings. A capability to train against a dynamic thinking enemy could enhance home station and pre-deployment small unit training. With the use of distributed virtual environments, this training could occur not only in post simulation facilities but just about anywhere a high speed internet connection exists. Further, with appropriate security measures, any centralized facility could support training of deployed units at remote locations. Finally, as more databases become available, mission rehearsal exercises may be possible down to the patrol level.

ACKNOWLEDGEMENTS

The research effort described in this paper was sponsored through a Small Business Innovative Research effort funded by the Office of the Secretary of Defense and directed by the Office of Naval Research.

REFERENCES

Andre, D., & Russell, S. (2002). *State abstraction for programmable reinforcement learning agents.* Paper presented at the AAAI-02, Edmonton, Alberta.

Dahlbom, A., & Niklasson, L. (2006). Goal-directed hierarchical dynamic scripting for RTS games. Paper presented at the Second Artificial Intelligence in Interactive Digital Entertainment, Marina del Rey, California.

Fu, D., & Houlette, R. (2002). Putting AI in Entertainment: An AI Authoring Tool for Simulation and Games. *IEEE Intelligent Systems*(July-August), 81-84.

Ludwig, J., & Farley, A. (2008). *Using Hierarchical Dynamic Scripting to Create Adaptive Adversaries.*

Paper presented at the 2008 Conference on Behavior Representation in Modeling and Simulation, Providence, RI.

Ponsen, M., & Spronck, P. (2004). Improving adaptive game ai with evolutionary learning. Paper presented at the CGAIDE 2004 International Conference on Computer Games.

Purpose Driven Virtual Worlds for Everyone. From http://www.forterrainc.com/images/stories/pdf/OLIVE_Dec07_Final_Rev.pdf

Spronck, P., Ponsen, M., Sprinkhuizen-Kuyper, I., & Postma, E. (2006). Adaptive game AI with dynamic scripting. *Machine Learning, 63*(3), 217-248.

Wiggins, M. (1997). "Pyramid of Learning," Foundations of Cooperative Learning Southeastern Center for Cooperative Learning, Jacksonville, FL